

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 1 1 月 2 7 日
Date of Application:

出 願 番 号 特 願 2 0 0 3 - 3 9 6 6 1 9
Application Number:
[ST. 10/C]: [J P 2 0 0 3 - 3 9 6 6 1 9]

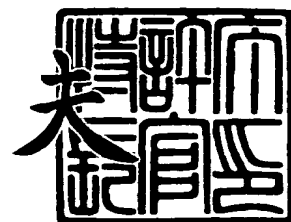
出 願 人 株式会社日立製作所
Applicant(s):

U.S. Appln. Filed 2-27-04
Inventor: T. Inoue et al
Mattingly Stanger & Malor
Docket NIT-413

2 0 0 4 年 1 月 2 8 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康



出証番号 出証特 2 0 0 4 - 3 0 0 3 6 8 9

【書類名】 特許願
【整理番号】 NT03P0836
【提出日】 平成15年11月27日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 12/501
G06F 13/12
G06F 13/38

【発明者】
【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
システム開発研究所内
【氏名】 井上 太郎

【発明者】
【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
システム開発研究所内
【氏名】 牧 晋広

【発明者】
【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所
ソフトウェア事業部内
【氏名】 小山田 健一

【特許出願人】
【識別番号】 000005108
【氏名又は名称】 株式会社日立製作所

【代理人】
【識別番号】 100068504
【弁理士】
【氏名又は名称】 小川 勝男
【電話番号】 03-3537-1621

【選任した代理人】
【識別番号】 100086656
【弁理士】
【氏名又は名称】 田中 恭助
【電話番号】 03-3537-1621

【選任した代理人】
【識別番号】 100094352
【弁理士】
【氏名又は名称】 佐々木 孝
【電話番号】 03-3537-1621

【手数料の表示】
【予納台帳番号】 081423
【納付金額】 21,000円

【提出物件の目録】
【物件名】 特許請求の範囲 1
【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

一つ以上の運用サイトと記憶装置サブシステムを有するサイトとがネットワークで接続された計算機システムにおいて、

前記運用サイトは、ホスト上で動作するコピー管理プログラムと、コピー管理情報とを備え、

前記コピー管理情報は、前記コピー管理プログラムによりリモートコピーをおこなう際の状態情報を含み、

しかも、この計算機システム内の記憶装置サブシステム内の記憶装置に格納されて、リモートコピーの各処理ごとに更新されることを特徴とする計算機システム。

【請求項 2】

運用サイトに障害があったときに、

運用可能な他のサイトのホスト上で動作するコピー管理プログラムにより、各サイトの記憶装置サブシステム内の記憶装置に格納されているコピー管理情報を参照して、

前記障害の直前まで進行したコピー処理の直後からコピーを実行することを特徴とする請求項 1 記載の計算機システム。

【請求項 3】

前記記憶装置サブシステムは、記憶装置サブシステム間で記憶装置内のデータをコピーすることができ、かつ、記憶装置サブシステム内で記憶装置内のデータをコピーすることができることを特徴とする請求項 1 記載の計算機システム。

【請求項 4】

前記コピー管理情報は、

サブシステム間コピーの状態情報、サブシステム内コピーの状態情報、および、このコピー管理情報が記録された時刻のタイムスタンプ情報が含まれることを特徴とする請求項 1 記載の計算機システム。

【請求項 5】

前記コピー管理情報は、この計算機システム内の全てのサイトの記憶装置サブシステム内の記憶装置に格納されることを特徴とする請求項 1 記載の計算機システム。

【請求項 6】

一つ以上の運用サイトと記憶装置サブシステムを有するサイトとがネットワークで接続された計算機システムの障害復旧方法において、

前記運用サイトは、ホスト上で動作するコピー管理プログラムにより、各サイトに運用サイトの記憶装置サブシステムの記憶装置内のデータを、他のサイトの記憶装置サブシステム内の記憶装置にリモートコピーするステップと、

前記運用サイトは、コピー処理ごとに、リモートコピーの状態を運用サイトの記憶装置サブシステム内の記憶装置に格納されたコピー管理情報に反映するステップと、

他のサイトは、コピー処理ごとに、リモートコピーの状態をそのサイトの記憶装置サブシステム内の記憶装置に格納されたコピー管理情報に反映するステップとを有することを特徴とする計算機システムの障害復旧方法。

【請求項 7】

運用サイトに障害があったときに、

運用可能な他のサイトのホスト上で動作するコピー管理プログラムにより、各サイトの記憶装置サブシステム内の記憶装置に格納されているコピー管理情報を参照して、

前記障害の直前まで進行したコピー処理の直後からコピーを実行する処理をおこなうことを特徴とする請求項 6 記載の計算機システムの障害復旧方法。

【請求項 8】

前記運用サイトからリモートコピーするステップは、

記憶装置サブシステム間で記憶装置内のデータをコピーする処理と、記憶装置サブシステム内で記憶装置内のデータをコピーする処理とからなり、

前記コピー管理情報にその各々のコピー処理について状態情報を保持することを特徴と

●
する請求項 6 記載の計算機システムの障害復旧方法。

【書類名】明細書**【発明の名称】**障害復旧機能を有する計算機システム、および、障害復旧方法**【技術分野】****【0001】**

本発明は、障害復旧機能を有する計算機システムに係り、特に、記憶装置サブシステム間でデータをリモートコピーすることにより、災害時におけるデータ保全をおこなうディザスタリカバリシステムに用いて好適な障害復旧機能を有する計算機システムに関する。

【背景技術】**【0002】**

システムの大規模化とデータの安全性を高めるために、通常の運用がおこなわれるサイトとバックアップサイトの2ヶ所のデータセンターとからなるシステム構成（以下、この構成を「2DC構成」と呼ぶ）が一般化されつつある。このような構成におけるディザスタリカバリシステム（災害復旧システム）については、例えば、非特許文献1に開示された技術が知られている。

【0003】

【非特許文献1】 EMC社ホームページ：http://japan.emc.com/local/ja/JP/products/product_pdfs/srdf/srdf.pdf

【発明の開示】**【発明が解決しようとする課題】****【0004】**

運用がおこなわれるサイトとバックアップサイトの2ヶ所のデータセンターとから構成される2DC構成に対して、3ヶ所のデータセンターからなる構成（以下、この構成を「3DC構成」と呼ぶ）のディザスタリカバリシステムを考える。以下では、通常の運用がおこなわれるサイトを「プロダクションサイト」、プロダクションサイトの近隣にあるサイトを「ローカルサイト」、プロダクションサイトの遠隔地にあるサイトを「リモートサイト」と呼ぶ。

【0005】

この3DC構成のディザスタリカバリシステムでは、プロダクションサイトのストレージのボリュームは、ローカルサイトのストレージのボリュームへリモートコピーされ、ローカルサイトのストレージのボリュームは、リモートサイトのストレージのボリュームへリモートコピーされることにより、データが伝播されていく。このとき、プロダクションサイトのホストで動作するディザスタリカバリ管理プログラム（以下、これを「ディザスタリカバリ管理プログラムP」と呼ぶ）が、上記の一連のリモートコピーの制御をおこなう。

【0006】

ところで、従来のリモートコピー技術では、上記通常の運用がおこなわれるプロダクションサイトが被災した際に、当該サイトのホスト上で動作するディザスタリカバリ管理プログラムが管理するディザスタリカバリの実行に必要な管理情報を、他のローカルサイトやリモートサイトに引き継ぐことができず、災害時のデータの速やかな復旧をおこなうことができないと言う問題点があった。

【0007】

本発明は、上記問題点を解決するためになされたもので、その目的は、通常の運用がおこなわれるプロダクションサイトが被災した際に、他の障害がないサイトからディザスタリカバリ管理プログラムを実行させ、リモートコピーを再開することにより、災害時のデータの速やかな復旧をおこなうことができる障害復旧機能を有する計算機システムを提供することにある。

【課題を解決するための手段】**【0008】**

本発明では、プロダクションサイトは、コピー管理テーブルにリモートコピーのコピーの処理ごとに、コピー状態を記録する。そして、各サイトにもコピー管理テーブルを保持

しておき、同期してそのコピー状態を更新していく。

【0009】

プロダクションサイトが被災した場合には、リモートサイトのホストで動作するディザスタリカバリ管理プログラム（以下、これを「ディザスタリカバリ管理プログラムR」と呼ぶ）が、災害のおこっていないサイトのコピー管理テーブルを参照し、災害の直後からコピー処理を再開する。

【0010】

このようにすれば、通常の運用サイトが被災する直前の最新のデータが運用可能な状態にある別サイトにコピーされるので、このデータを利用して運用可能な状態にある別サイトにおいて通常運用サイトで動作していたアプリケーションプログラムを立ち上げることにより業務を再開できるようになる。

【0011】

以上では、3ヶ所のデータセンターからなる3DC構成のディザスタリカバリシステムの場合について述べたが、4ヶ所以上のデータセンターからなるディザスタリカバリシステムの場合でも同様である。

【発明の効果】

【0012】

本発明によれば、通常の運用がおこなわれるプロダクションサイトが被災した際に、他の障害がないサイトからディザスタリカバリ管理プログラムを実行させ、リモートコピーを再開することにより、災害時のデータの速やかな復旧をおこなうことができる。

【発明を実施するための最良の形態】

【0013】

以下、本発明に係る一実施形態を、図1ないし図12を用いて説明する。

【0014】

先ず、図1を用いて本発明の障害復旧機能を有する計算機システムの構成について説明する。

【0015】

図1は、本発明の障害復旧機能を有する計算機システムの構成図である。

【0016】

本システムは、ホスト1a、ホスト1c、記憶装置サブシステム2a、記憶装置サブシステム2b、記憶装置サブシステム2c、および、管理用計算機8を有する。これらの装置はネットワーク9により相互に接続されている。本システムでは、ホスト1aおよび記憶装置サブシステム2aをまとめて、プロダクションサイト3と呼び、記憶装置サブシステム2bを、ローカルサイト4と呼び、ホスト1cおよび記憶装置サブシステム2cをまとめて、リモートサイト5と呼ぶ。ローカルサイトには、ホストがないものとする。プロダクションサイト3とローカルサイト4は、通常地理的には近隣地（数十kmの距離）であり、ローカルサイト4とリモートサイト5は、通常地理的には遠隔地（数百kmの距離）であることを想定している。プロダクションサイト3は、通常運用が実行されるサイトであり、ローカルサイト4およびリモートサイト5は、プロダクションサイト3のバックアップをおこなうためのサイトである。本実施形態では、上記のように3ヶ所のサイトの場合を示しているが、4ヶ所以上のサイトの場合でも同様に本発明を適用できる。

【0017】

ホスト1および管理用計算機8は、計算機であり、CPU101、メモリ102、ディスプレイ103、キーボード104、コントローラ105、データインタフェース106、ネットワークインタフェース107、および、ディスク装置108を有する。

【0018】

記憶装置サブシステム2は、データを格納する記憶装置であり、ポート201、ディスクコントローラ202、制御メモリ203、プロセッサ204、キャッシュメモリ205、ディスク装置211を有する。

【0019】

ディスク装置 108 および 211 は、論理的な記憶装置であり、実際には複数の物理的記憶装置が 1 つの論理的な記憶装置を構成していてもよい。この場合、複数の物理的記憶装置がディスクアレイを構成していてもよい。ここで物理的記憶装置とは、例えば、ハードディスクドライブや DVD といった物理的な記憶装置である。

本システムにおけるリモートコピーの動作の一例を説明する。プライマリサイト 3 のディスク装置 211 a のデータは、ローカルサイト 4 のディスク装置 211 b にリモートコピーされる。ディスク装置 211 b のデータは、同じローカルサイト内のディスク装置 211 c にコピーされ、ディスク装置 211 c のデータは、リモートサイトのディスク装置 211 d にリモートコピーされる。ディスク装置 211 d のデータは、同じリモートサイト内のディスク装置 211 e にコピーされる。このように、通常運用をおこなうプロダクションサイト 3 のデータは、順次ローカルサイト 4、リモートサイト 5 へとコピーされる。

【0020】

次に、図 2 ないし図 5 を用いて本システムでのプログラムとデータのメモリおよびディスク装置内の格納状況について説明する。

【0021】

図 2 は、プロダクションサイト 3 のホスト 1 a のディスク装置 108 a の格納内容を示す図である。

【0022】

図 3 は、リモートサイト 5 のホスト 1 c のディスク装置 108 c の格納内容を示す図である。

【0023】

図 4 は、制御メモリ 203 の格納内容の詳細を示す図である。

【0024】

図 5 は、管理用計算機 8 のディスク装置 108 d の格納内容を示す図である。

【0025】

ディスク装置 108 a には、図 2 に示されるようにディザスタリカバリ管理プログラム P120 が格納される。このプログラムは、メモリ 102 a に読み出され、CPU 101 a で実行される。ディザスタリカバリ管理プログラム P120 は、前記のプロダクションサイト 3 からローカルサイト 4 を経てリモートサイト 5 へ至るリモートコピーの制御をおこなう。

【0026】

ディスク装置 108 c には、図 3 に示されるようにディザスタリカバリ管理プログラム R121 が格納される。ディザスタリカバリ管理プログラム R121 は、メモリ 102 c に読み出され、CPU 101 c で実行される。ディザスタリカバリ管理プログラム R121 は、プロダクションサイト 3 が被災した場合、あるいは、プロダクションサイト 3 およびローカルサイト 4 の両者が被災した場合にリカバリ処理をおこなう。

制御メモリ 203 には、図 4 に示されるようにサブシステム間ペア定義テーブル 300、サブシステム内ペア定義テーブル 400、および、ボリューム定義テーブル 500 が格納される。これらの後に詳細に説明する。

【0027】

ディスク装置 108 d には、図 5 に示されるようにボリューム定義プログラム 122 が格納される。ボリューム定義プログラム 122 は、メモリ 102 d に読み出され、CPU 101 d で実行される。ボリューム定義プログラム 122 の詳細は、後ほど説明する。

【0028】

次に、図 6 ないし図 9 を用いて本システムで用いられテーブルのデータ構造について説明する。

【0029】

図 6 は、サブシステム間ペア定義テーブル 300 の構成の例を示す図である。

【0030】

図7は、サブシステム内ペア定義テーブル400の構成の例を示す図である。

【0031】

図8は、ボリューム定義テーブル500の構成の例を示す図である。

【0032】

図9は、管理情報格納用ボリュームに格納されるコピー管理テーブル600の内容を示す図である。

【0033】

サブシステム間ペア定義テーブル300は、記憶装置サブシステム2の制御メモリ203に格納される。サブシステム間ペア定義テーブル300は、図6に示されるようにグループ名フィールド301、ペア名フィールド302、正ポートフィールド303、正論理ユニットフィールド304、副ポートフィールド305、副論理ユニットフィールド306、および、状態フィールド307を有する。

【0034】

グループ名フィールド301は、図6に示されるようにペア関係の集合であるグループの名称が登録されるフィールドである。ペア名フィールド302は、ペア関係の名称が登録されるフィールドである。正ポートフィールド303は、ペア関係を構成する正サイトの記憶装置サブシステムのポートの情報が登録されるフィールドである。正論理ユニットフィールド304は、ペア関係を構成する正サイトの記憶装置サブシステムの論理ユニット名が登録されるフィールドである。副ポートフィールド305は、ペア関係を構成する副サイトの記憶装置サブシステムのポートの情報が登録されるフィールドである。副論理ユニットフィールド306は、ペア関係を構成する副サイトの記憶装置サブシステムの論理ユニット名が登録されるフィールドである。状態フィールド307は、ペア関係の状態を示す情報が登録されるフィールドである。

【0035】

ここで、論理ユニット（以下「LU」）は、ディスク装置211で構成される記憶領域を管理する際の単位である。また、ボリュームは、LUへのアクセスに使用されるポートの名称とボリュームを構成するLUの名称との組み合わせにより一意に識別される。以下、ボリュームを表記する際に、ポートの名称およびLUの名称を用いてボリューム（ポート名称、LU名称）と表記する。なお、本実施形態では、LU一つが一つのボリュームに対応するとして説明するが、複数のLUが一つのボリュームを構成しても良い。

【0036】

この図6の例では、図1の構成におけるサブシステム間のペア構成を示している。レコード310aでは、グループG1に属するペアP1があり、P1を構成する正サイトのボリュームがボリューム（ポート201b、LU0）、P1を構成する副サイトのボリュームがボリューム（ポート201d、LU0）、および、状態がACTIVEであることが示されている。ここで、ACTIVEとは、レコード310aで示されたペア間でデータの複製処理、すなわち正副サイトのデータの内容を一致させる処理がおこなわれている最中であることを示している。

【0037】

次に、サブシステム内ペア定義テーブル400は、記憶装置サブシステム2の制御メモリ203に格納される。サブシステム内ペア定義テーブル400は、図7に示されるようにペア名フィールド402、正論理ユニットフィールド403、副論理ユニットフィールド404、状態フィールド405を有する。

【0038】

ペア名フィールド402は、ペア関係の名称が登録されるフィールドである。正論理ユニットフィールド403は、ペア関係を構成する正論理ユニット名が登録されるフィールドである。副論理ユニットフィールド404は、ペア関係を構成する副論理ユニット名が登録されるフィールドである。状態フィールド405はペア関係の状態を示す情報が登録されるフィールドである。

【0039】

図7の例では、図1の構成におけるサブシステム内のペア構成を示している。サブシステム内ペア定義テーブル400bは、制御メモリ203bに格納され、サブシステム内ペア定義テーブル400cは、制御メモリ203cに格納される。レコード410aでは、ペアP10があり、P10を構成する正ボリュームがボリューム（ポート201d, LU0）、P1を構成する副ボリュームがボリューム（ポート201d, LU1）、および、状態がACTIVEであることが示されている。

【0040】

ここで、ACTIVEとは、レコード410aで示されたペア間でデータの複製処理、すなわち正副論理ユニットのデータの内容を一致させる処理がおこなわれている最中であることを示している。

【0041】

次に、ボリューム定義テーブル500は、記憶装置サブシステム2の制御メモリ203に格納される。ボリューム定義テーブル500は、図8に示されるようにポートフィールド501、LUフィールド502、ドライブフィールド503、サイズフィールド504、管理情報格納フィールド505を有する。

【0042】

ポートフィールド501は、ボリュームを識別するポートの情報が登録されるフィールドである。LUフィールド502は、ボリュームを構成するLUの情報が登録されるフィールドである。ドライブフィールド503は、ボリュームを構成するディスク装置211の情報が登録されるフィールドである。サイズフィールド504は、当該ボリュームの容量の情報が登録されるフィールドである。管理情報格納フィールド505は、管理情報を格納するか否かを示すフィールドである。

【0043】

ボリューム定義テーブル500は、ユーザがボリューム定義プログラム122を用いて予め設定する。ここで、ボリューム定義プログラム122とは、ネットワーク3に接続される管理用計算機8で実行されるプログラムである。ユーザは、管理用計算機8を介して、記憶装置サブシステム2に設定すべきボリュームの情報を記憶装置サブシステム2に送付する。記憶装置サブシステム2は、送付された情報をボリューム管理テーブル500に登録する。

【0044】

図8の例では、図1の構成におけるボリュームを示している。ボリューム定義テーブル500aは制御メモリ203aに格納され、ボリューム定義テーブル500bは制御メモリ203bに格納され、ボリューム定義テーブル500cは制御メモリ203cに格納される。レコード510bでは、ボリューム（ポート201b, LU1）はディスク装置211fから構成され、その容量は10GB（ギガバイト）であり、管理情報格納がYESであることが示されている。ここで、YESとは管理情報を格納することを示している。すなわち、管理情報格納用ボリュームであることを示している（以下、管理情報格納用ボリュームを管理情報格納用ディスク装置と呼ぶ場合もある）。

【0045】

次に、図9に示されるコピー管理テーブル600は、サブシステム間コピーおよびサブシステム内コピーの状況を管理する情報を格納するためのテーブルである。コピー管理テーブル600は、あらかじめ定められたファイル名を有するファイルに格納されており、ディザスタリカバリ管理プログラムP120が、プライマリサイト3の管理情報格納用ディスク装置（ディスク装置211a）の当該ファイルに随時上書きして更新する。

【0046】

そして、この内容は、サブシステム間コピー機能およびサブシステム内コピー機能により、順次ディスク装置211b、ディスク装置211c、ディスク装置211d、ディスク装置211eと伝えられていく。

【0047】

記憶装置サブシステム2aから記憶装置サブシステム2bへのサブシステム間コピーの

状態 601 には、実行状態および実行（凍結）状態が設定される。記憶装置サブシステム 2b のサブシステム内コピーの状態 602 には、実行状態および切り離し状態が設定される。記憶装置サブシステム 2b から記憶装置サブシステム 2c へのサブシステム間コピーの状態 603 には、実行状態および中断状態が設定される。記憶装置サブシステム 2c のサブシステム内コピーの状態 604 には、実行状態および切り離し状態が設定される。タイムスタンプ 605 には、当該レコードが記録された時刻情報が設定される。

【0048】

次に、図 10 ないし図 12 を用いて本発明に係る計算機システムの障害復旧の処理手順について説明する。

【0049】

まず、図 10 によりボリューム定義プログラム 122 の処理について説明する。

【0050】

図 10 は、ボリューム定義プログラム 122 の処理を示すフローチャートである。

【0051】

ボリューム定義プログラム 122 を起動されると（ステップ 1000）、ユーザは、ディスプレイ 103d およびキーボード 104d により、ボリューム定義をおこなう記憶装置サブシステム 2 の情報、ボリュームを識別するポートの情報、ボリュームを構成する LU の情報、ボリュームを構成するディスク装置 211 の情報、当該ボリュームの容量の情報、および管理情報を格納するか否かを示す情報を設定する（ステップ 1005）。

【0052】

そして、設定された定義情報を指定された記憶装置サブシステム 2 に送付して（ステップ 1010）、処理を終了する（ステップ 1015）。定義情報が送付された記憶装置サブシステム 2 では、定義情報をボリューム定義テーブル 500 に設定する。

【0053】

次に、図 11 を用いてディザスタリカバリ管理プログラム P120 の処理を説明する。

【0054】

図 11 は、ディザスタリカバリ管理プログラム P120 の処理を示すフローチャートである。

【0055】

ディザスタリカバリ管理プログラム P120 は、メモリ 102a に読み出され、CPU 101a で実行される。ディザスタリカバリ管理プログラム P120 は、プロダクションサイト 3 からローカルサイト 4 を経てリモートサイト 5 へ至るリモートコピーの制御をおこなう。

【0056】

以下では、記憶装置サブシステム 2a から記憶装置サブシステム 2b へサブシステム間コピーをおこない、記憶装置サブシステム 2b のサブシステム内コピーをおこない、記憶装置サブシステム 2b から記憶装置サブシステム 2c へサブシステム間コピーをおこない、記憶装置サブシステム 2c のサブシステム内コピーをおこなう制御を説明する。すなわち、この過程で、管理情報格納用ボリュームであるディスク装置 211f からディスク装置 211g へサブシステム間コピーし、ディスク装置 211g からディスク装置 211h へサブシステム内コピーし、ディスク装置 211h からディスク装置 211i へサブシステム間コピーし、ディスク装置 211i からディスク装置 211j へサブシステム内コピーがおこなわれる。

【0057】

そして、管理情報格納用ディスク装置ではないディスク装置 211a からディスク装置 211b へサブシステム間コピーし、ディスク装置 211b からディスク装置 211c へサブシステム内コピーし、ディスク装置 211c からディスク装置 211d へサブシステム間コピーし、ディスク装置 211d からディスク装置 211e へサブシステム内コピーがおこなわれる。なお、以下でのサブシステム間コピーおよびサブシステム内コピーの指示は、ディスクコントローラ 202 への命令によってなされる。

【0058】

ディザスタリカバリ管理プログラム P120 が起動されると（ステップ 1100）、記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの実行を指示し、その他のサブシステム間コピーは中断を指示し、サブシステム内コピーは切り離しを指示する（ステップ 1105）。この時点でのサブシステム間コピーおよびサブシステム内コピーの状態（記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの状態 601：実行状態、記憶装置サブシステム 2b のサブシステム内コピーの状態 602：切り離し状態、記憶装置サブシステム 2b から記憶装置サブシステム 2c へのサブシステム間コピーの状態 603：中断状態、記憶装置サブシステム 2c のサブシステム内コピーの状態 604：切り離し状態）、および、タイムスタンプ 605 をコピー管理テーブル 600 に設定する（ステップ 1107）。

【0059】

次に、記憶装置サブシステム 2b のサブシステム内コピーの再同期、および記憶装置サブシステム 2c のサブシステム内コピーの再同期の実行を指示する（ステップ 1110）。この時点でのサブシステム間コピーおよびサブシステム内コピーの状態（記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの状態 601：実行状態、記憶装置サブシステム 2b のサブシステム内コピーの状態 602：実行状態、記憶装置サブシステム 2b から記憶装置サブシステム 2c へのサブシステム間コピーの状態 603：中断状態、記憶装置サブシステム 2c のサブシステム内コピーの状態 604：実行状態）、および、タイムスタンプ 605 をコピー管理テーブル 600 に設定する（ステップ 1112）。

【0060】

次に、記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの実行の凍結を指示する（ステップ 1115）。この時点でのサブシステム間コピーおよびサブシステム内コピーの状態（記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの状態 601：実行（凍結）状態、記憶装置サブシステム 2b のサブシステム内コピーの状態 602：実行状態、記憶装置サブシステム 2b から記憶装置サブシステム 2c へのサブシステム間コピーの状態 603：中断状態、記憶装置サブシステム 2c のサブシステム内コピーの状態 604：実行状態）、および、タイムスタンプ 605 をコピー管理テーブル 600 に設定する（ステップ 1117）。

【0061】

次に、記憶装置サブシステム 2b のサブシステム内コピーの切り離し、および記憶装置サブシステム 2c のサブシステム内コピーの切り離しを指示する。（ステップ 1120）。この時点でのサブシステム間コピーおよびサブシステム内コピーの状態（記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの状態 601：実行（凍結）状態、記憶装置サブシステム 2b のサブシステム内コピーの状態 602：切り離し状態、記憶装置サブシステム 2b から記憶装置サブシステム 2c へのサブシステム間コピーの状態 603：中断状態、記憶装置サブシステム 2c のサブシステム内コピーの状態 604：切り離し状態）、および、タイムスタンプ 605 をコピー管理テーブル 600 に設定する（ステップ 1122）。

【0062】

その後、記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの実行の再開を指示する（ステップ 1125）。この時点でのサブシステム間コピーおよびサブシステム内コピーの状態（記憶装置サブシステム 2a から記憶装置サブシステム 2b へのサブシステム間コピーの状態 601：実行状態、記憶装置サブシステム 2b のサブシステム内コピーの状態 602：切り離し状態、記憶装置サブシステム 2b から記憶装置サブシステム 2c へのサブシステム間コピーの状態 603：中断状態、記憶装置サブシステム 2c のサブシステム内コピーの状態 604：切り離し状態）、および、タイムスタンプ 605 をコピー管理テーブル 600 に設定する（ステップ 1127）。

【0063】

そして、記憶装置サブシステム 2 b から記憶装置サブシステム 2 c へのサブシステム間コピーの再同期の実行を指示する（ステップ 1130）。この時点でのサブシステム間コピーおよびサブシステム内コピーの状態（記憶装置サブシステム 2 a から記憶装置サブシステム 2 b へのサブシステム間コピーの状態 601：実行状態、記憶装置サブシステム 2 b のサブシステム内コピーの状態 602：切り離し状態、記憶装置サブシステム 2 b から記憶装置サブシステム 2 c へのサブシステム間コピーの状態 603：実行状態、記憶装置サブシステム 2 c のサブシステム内コピーの状態 604：切り離し状態）、および、タイムスタンプ 605 をコピー管理テーブル 600 に設定する（ステップ 1132）。

【0064】

次に、記憶装置サブシステム 2 b から記憶装置サブシステム 2 c へのサブシステム間コピーの中断を指示する（ステップ 1135）。この時点でのサブシステム間コピーおよびサブシステム内コピーの状態（記憶装置サブシステム 2 a から記憶装置サブシステム 2 b へのサブシステム間コピーの状態 601：実行状態、記憶装置サブシステム 2 b のサブシステム内コピーの状態 602：切り離し状態、記憶装置サブシステム 2 b から記憶装置サブシステム 2 c へのサブシステム間コピーの状態 603：中断状態、記憶装置サブシステム 2 c のサブシステム内コピーの状態 604：切り離し状態）、および、タイムスタンプ 605 をコピー管理テーブル 600 に設定する（ステップ 1137）。

【0065】

これでステップ 1105 の実行直後の状態と同じになるので、ステップ 1110 以降の処理を繰り返し実行する。以上の処理により、プロダクションサイト 3 のディスク装置 211 a およびディスク装置 211 f の内容は、ローカルサイト 4 を経てリモートサイト 5 のディスク装置 211 e およびディスク装置 211 j にコピーされる。

【0066】

ここで、ステップ 1107、ステップ 1112、ステップ 1117、ステップ 1122、ステップ 1127、ステップ 1132、ステップ 1137 で、コピー管理テーブルが更新されたときには、同期して他のサイトのコピー管理テーブルも更新されることを注意すべきである。

【0067】

ディザスタリカバリ管理プログラム R121 は、プロダクションサイト 3 が被災した場合、あるいは、プロダクションサイト 3 とローカルサイト 4 の両者が被災した場合に、リカバリ処理をおこなうプログラムで、リモートサイト 5 において、メモリ 102 c に読み出され、CPU 101 c で実行される。

【0068】

最後に、図 12 を用いてプロダクションサイト 3 のみが被災した場合のディザスタリカバリ管理プログラム R121 の処理を説明する。

【0069】

図 12 は、プロダクションサイト 3 のみが被災した場合のディザスタリカバリ管理プログラム R121 の処理を示すフローチャートである。

【0070】

先ず、災害の発生に伴いディザスタリカバリ管理プログラム R121 を起動する（ステップ 1200）。災害後に生き残っているすべての管理情報格納用ディスク装置からコピー管理テーブル 600 を収集する（ステップ 1210）。そして、収集したコピー管理テーブル 600 におけるタイムスタンプ 605 のタイムスタンプ値をチェックし、最新のタイムスタンプ値を有するコピー管理テーブル 600 におけるコピー状態情報（記憶装置サブシステム 2 a から記憶装置サブシステム 2 b へのサブシステム間コピーの状態 601、記憶装置サブシステム 2 b のサブシステム内コピーの状態 602、記憶装置サブシステム 2 b から記憶装置サブシステム 2 c へのサブシステム間コピーの状態 603、記憶装置サブシステム 2 c のサブシステム内コピーの状態 604）を基に、図 11 のフローチャートに示した処理のどこまで進行しているかを把握したうえで、図 11 のフローチャートに示した処理の続きを再開し（ステップ 1220）、最後のコピー処理を終えたときに処理を

終了する（ステップ1230）。この結果、リモートサイト5のディスク装置211eおよびディスク装置211jまで、データが伝播することになる。

プロダクションサイト3とローカルサイト4の両者が被災した場合には、生き残っているのはリモートサイト5のみなので、ディザスタリカバリ管理プログラムR121は、図11のステップ1110の処理のみを実行する。

【0071】

以上により、プロダクションサイト3が被災する直前の最新のデータがリモートサイト5に伝播したので、このデータを利用してリモートサイト5においてプロダクションサイト3で動作していたアプリケーションプログラムを立ち上げることにより、業務を再開できるようにする。

【図面の簡単な説明】

【0072】

【図1】本発明の障害復旧機能を有する計算機システムの構成図である。

【図2】プロダクションサイト3のホスト1aのディスク装置108aの格納内容を示す図である。

【図3】リモートサイト5のホスト1cのディスク装置108cの格納内容を示す図である。

【図4】制御メモリ203の格納内容の詳細を示す図である。

【図5】管理用計算機8のディスク装置108dの格納内容を示す図である。

【図6】サブシステム間ペア定義テーブル300の構成の例を示す図である。

【図7】サブシステム内ペア定義テーブル400の構成の例を示す図である。

【図8】ボリューム定義テーブル500の構成の例を示す図である。

【図9】管理情報格納用ボリュームに格納されるコピー管理テーブル600の内容を示す図である。

【図10】ボリューム定義プログラム122の処理を示すフローチャートである。

【図11】ディザスタリカバリ管理プログラムP120の処理を示すフローチャートである。

【図12】プロダクションサイト3のみが被災した場合のディザスタリカバリ管理プログラムR121の処理を示すフローチャートである。

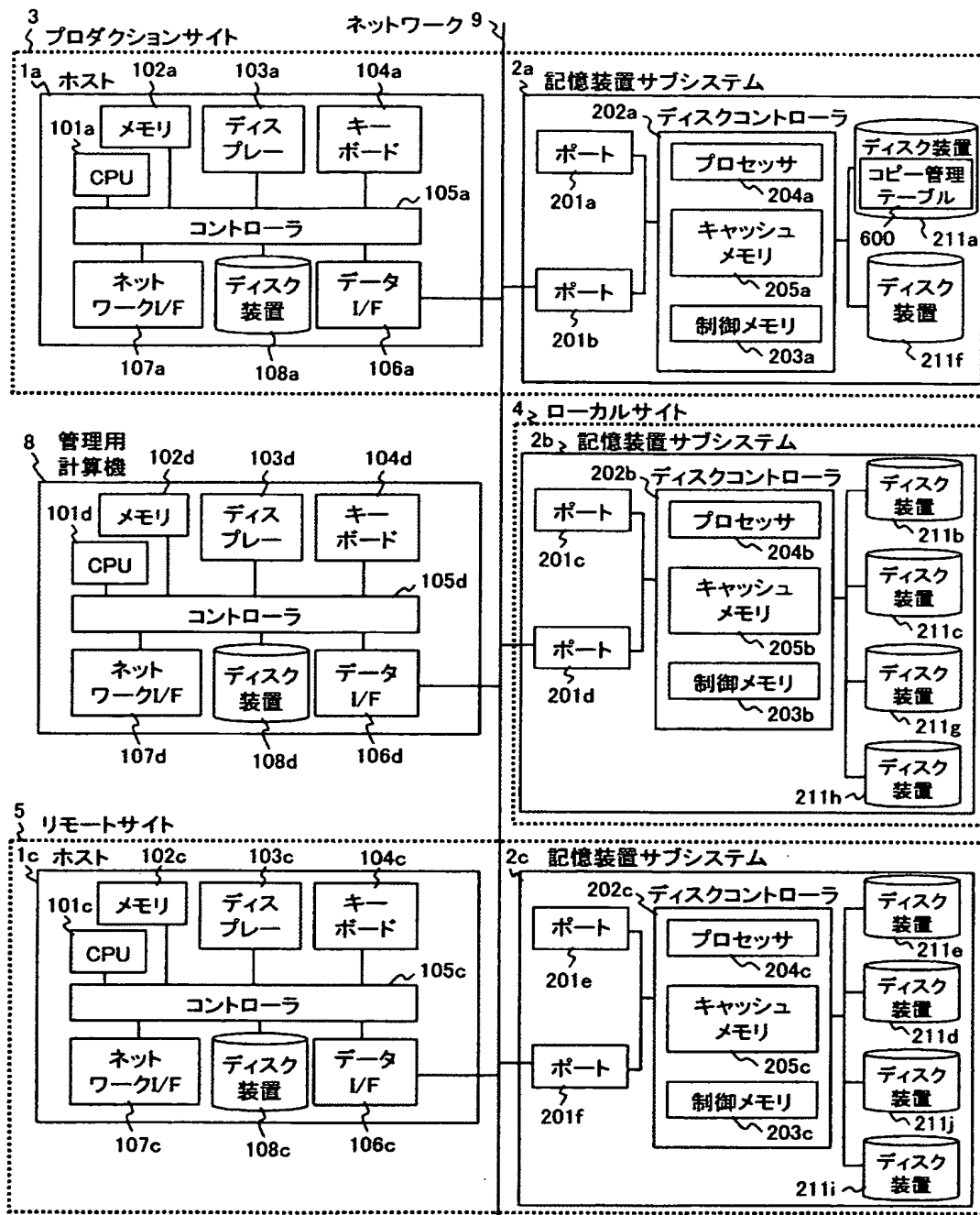
【符号の説明】

【0073】

1…ホスト、2…記憶装置サブシステム、3…プロダクションサイト、4…ローカルサイト、5…リモートサイト、8…管理用計算機、101…CPU、102…メモリ、103…ディスプレイ、104…キーボード、105…コントローラ、106…データインタフェース、107…ネットワークインタフェース、108…ディスク装置、120…ディザスタリカバリ管理プログラムP、121…ディザスタリカバリ管理プログラムR、122…ボリューム定義プログラム、201…ポート、202…ディスクコントローラ、203…制御メモリ、204…プロセッサ、205…キャッシュメモリ、211…ディスク装置、300…サブシステム間ペア定義テーブル、400…サブシステム内ペア定義テーブル、500…ボリューム定義テーブル、600…コピー管理テーブル。

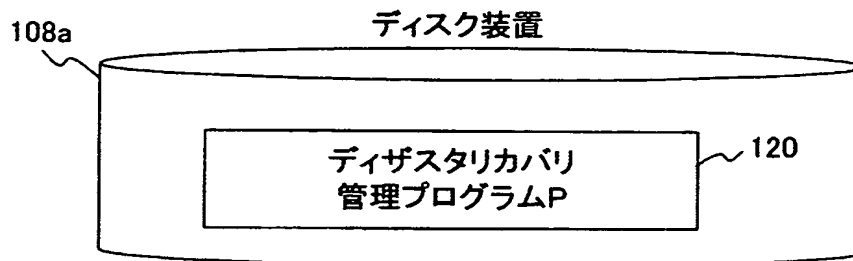
【書類名】 図面
【図 1】

図 1



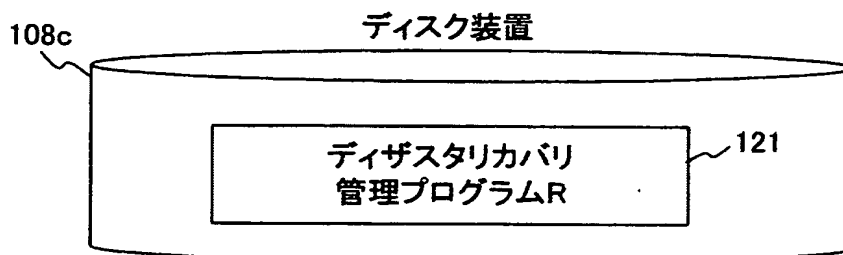
【図 2】

図 2



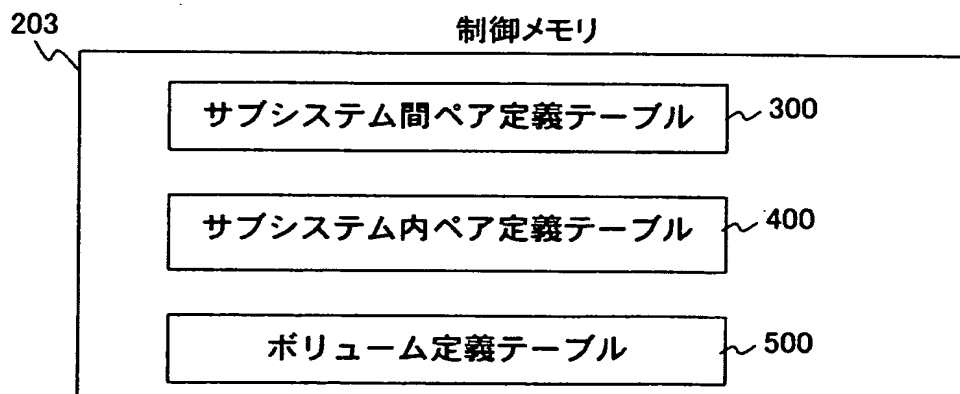
【図 3】

図 3



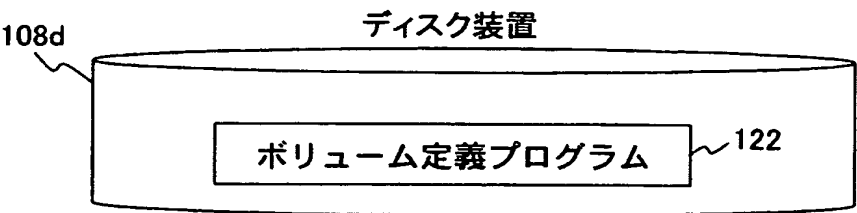
【図 4】

図 4



【図 5】

図 5



【図 6】

図 6

サブシステム間ペア定義テーブル

300	301	302	303	304	305	306	307
	グループ名	ペア名	正ポート	正LU	副ポート	副LU	状態
310a	G1	P1	ポート201b	LU0	ポート201d	LU0	ACTIVE
310b	G1	P2	ポート201d	LU1	ポート201f	LU1	ACTIVE
	G1	P3	ポート201b	LU1	ポート201d	LU2	ACTIVE
	G1	P4	ポート201d	LU3	ポート201f	LU3	ACTIVE

【図 7】

図 7

サブシステム内ペア定義テーブル

400b	401	402	403	404	405
	グループ名	ペア名	正LU	副LU	状態
410a	G11	P10	LU0	LU1	ACTIVE
410b	G11	P11	LU2	LU3	ACTIVE

サブシステム内ペア定義テーブル

400c	401	402	403	404	405
	グループ名	ペア名	正LU	副LU	状態
410c	G12	P12	LU0	LU1	ACTIVE
410d	G12	P13	LU2	LU3	ACTIVE

【図 8】

図 8

ボリューム定義テーブル

500a	401	402	403	404	405
	ポート	LU	ドライブ	サイズ	管理情報格納
510a	ポート201b	LU0	ディスク211a	10GB	
510b	ポート201b	LU1	ディスク211f	10GB	YES

ボリューム定義テーブル

500b	401	402	403	404	405
	ポート	LU	ドライブ	サイズ	管理情報格納
510c	ポート201d	LU0	ディスク211b	10GB	
510d	ポート201d	LU1	ディスク211c	10GB	
510e	ポート201d	LU2	ディスク211g	10GB	YES
510f	ポート201d	LU3	ディスク211h	10GB	YES

ボリューム定義テーブル

500c	401	402	403	404	405
	ポート	LU	ドライブ	サイズ	管理情報格納
510g	ポート201f	LU0	ディスク211e	10GB	
510h	ポート201f	LU1	ディスク211d	10GB	
510i	ポート201f	LU2	ディスク211j	10GB	YES
510j	ポート201f	LU3	ディスク211i	10GB	YES

【図 9】

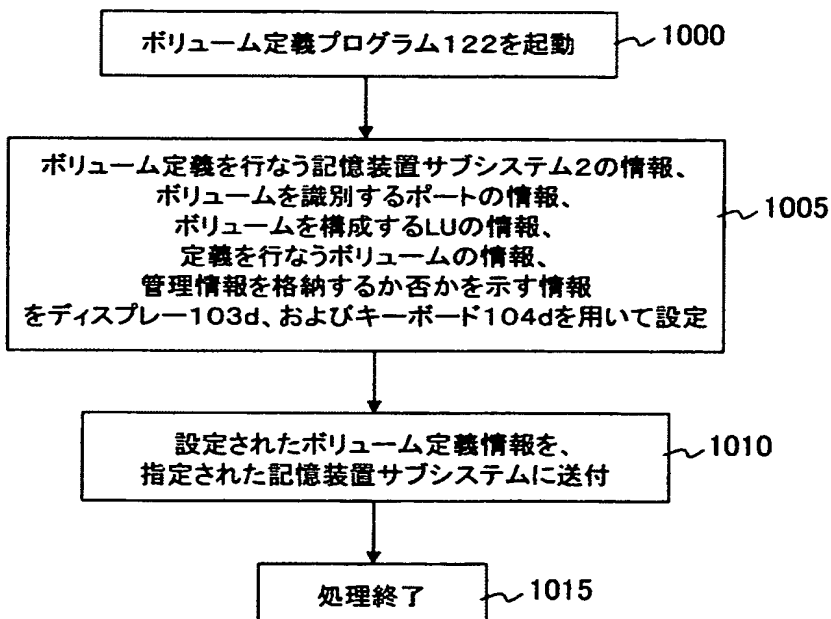
図 9

コピー管理テーブル

600	601	602	603	604	605
記憶装置サブシステム2aから記憶装置サブシステム2bへのサブシステム間コピーの状態	記憶装置サブシステム2bのサブシステム内コピーの状態	記憶装置サブシステム2bから記憶装置サブシステム2cへのサブシステム間コピーの状態	記憶装置サブシステム2cのサブシステム内コピーの状態	タイムスタンプ	
610a 実行状態	切り離し状態	中断状態	切り離し状態	1526874526	

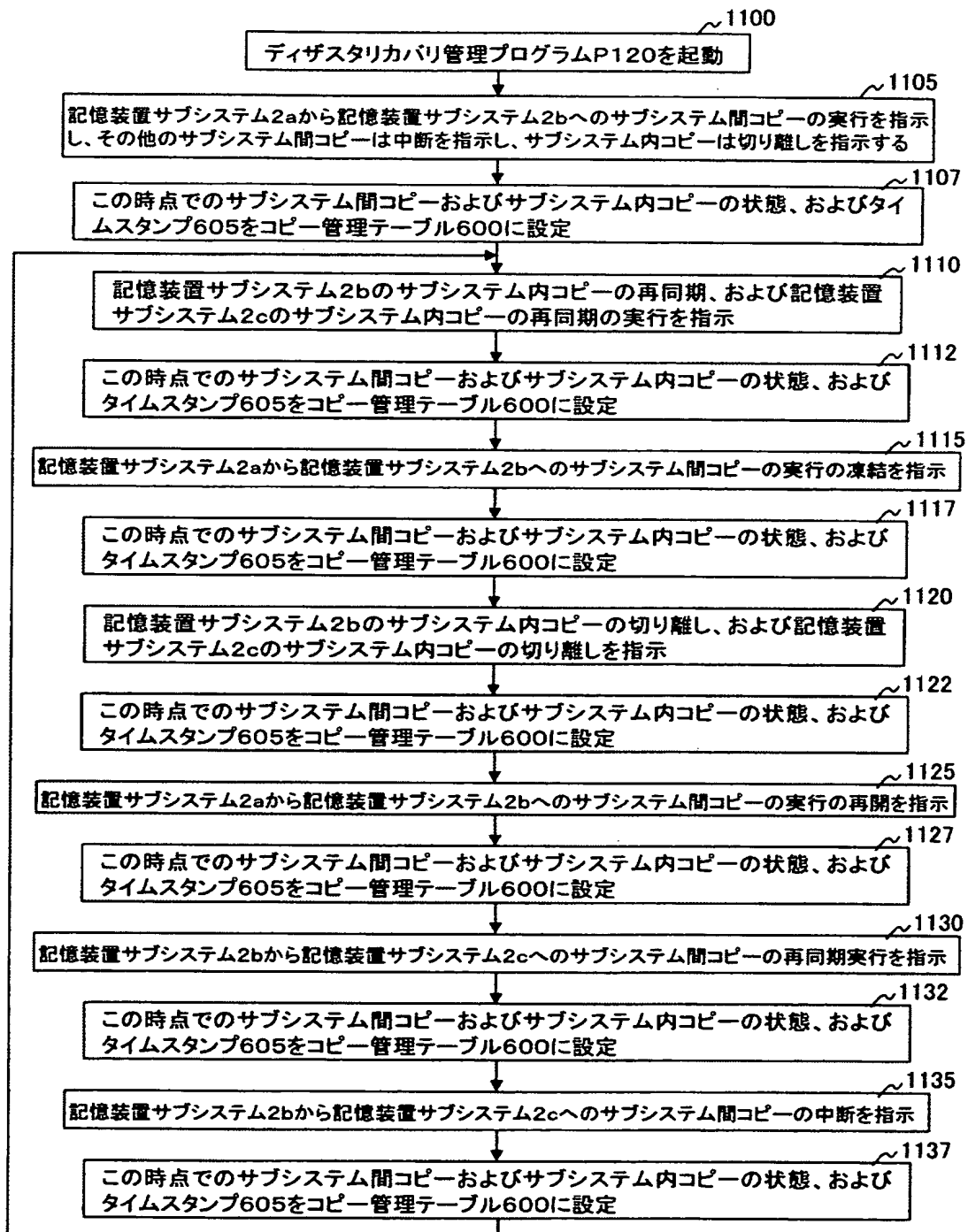
【図 10】

図 10



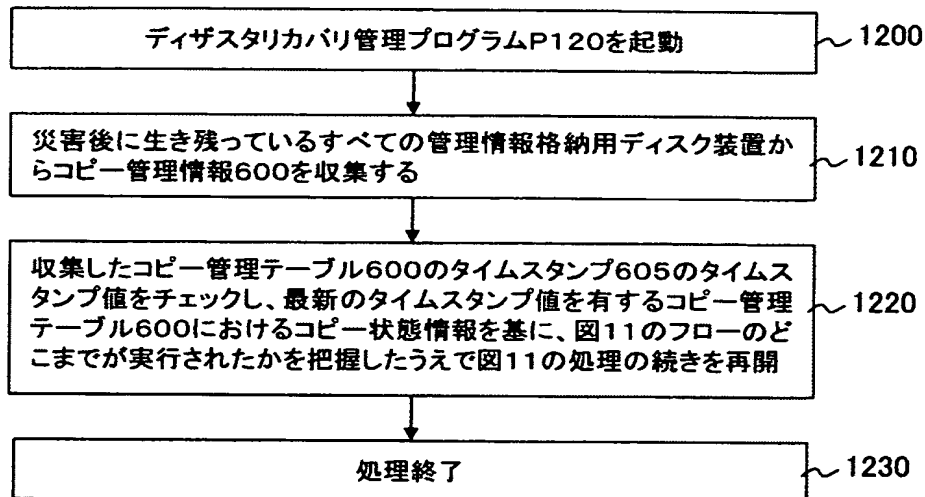
【図 11】

図 11



【図 12】

図 12



【書類名】 要約書**【要約】**

【課題】 アプリケーションの運用がおこなわれるプロダクションサイトが被災した際に、他の障害がないサイトからディザスタリカバリ管理プログラムを実行させ、リモートコピーを再開することにより、災害時のデータの速やかな復旧をおこなう。

【解決手段】 運用サイトは、リモートコピーをおこなう際の状態情報を保持するコピー管理情報を備える。そして、この計算機システム内の全ての記憶装置サブシステム内の記憶装置にコピー管理情報を格納しておき、リモートコピーの各処理ごとに更新する。運用サイトに障害があったときに、運用可能な他のサイトのコピー管理プログラムにより、各サイトの記憶装置サブシステム内の記憶装置に格納されているコピー管理情報を参照して、障害の直前まで進行したコピー処理の直後からコピーを実行する。

【選択図】 図 1

特願 2 0 0 3 - 3 9 6 6 1 9

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所